



TECHNICAL PAPER

# Introduction to Closed Captions

By Glenn Eguchi  
Senior Computer Scientist  
April 2015

© 2015 Adobe Systems Incorporated. All rights reserved.

If this whitepaper is distributed with software that includes an end user agreement, this guide, as well as the software described in it, is furnished under license and may be used or copied only in accordance with the terms of such license. Except as permitted by any such license, no part of this guide may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, recording, or otherwise, without the prior written permission of Adobe Systems Incorporated. Please note that the content in this guide is protected under copyright law even if it is not distributed with software that includes an end user license agreement.

The content of this guide is furnished for informational use only, is subject to change without notice, and should not be construed as a commitment by Adobe Systems Incorporated. Adobe Systems Incorporated assumes no responsibility or liability for any errors or inaccuracies that may appear in the informational content contained in this guide.

This article is intended for US audiences only.

Any references to company names in sample templates are for demonstration purposes only and are not intended to refer to any actual organization.

Adobe and the Adobe logo, and Adobe Primetime are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.

Adobe Systems Incorporated, 345 Park Avenue, San Jose, California 95110, USA. Notice to U.S. Government End Users. The Software and Documentation are "Commercial Items," as that term is defined at 48 C.F.R. §2.101, consisting of "Commercial Computer Software" and "Commercial Computer Software Documentation," as such terms are used in 48 C.F.R. §12.212 or 48 C.F.R. §227.7202, as applicable. Consistent with 48 C.F.R. §12.212 or 48 C.F.R. §§227.7202-1 through 227.7202-4, as applicable, the Commercial Computer Software and Commercial Computer Software Documentation are being licensed to U.S. Government end users (a) only as Commercial Items and (b) with only those rights as are granted to all other end users pursuant to the terms and conditions herein. Unpublished-rights reserved under the copyright laws of the United States. Adobe Systems Incorporated, 345 Park Avenue, San Jose, CA 95110-2704, USA. For U.S. Government End Users, Adobe agrees to comply with all applicable equal opportunity laws including, if appropriate, the provisions of Executive Order 11246, as amended, Section 402 of the Vietnam Era Veterans Readjustment Assistance Act of 1974 (38 USC 4212), and Section 503 of the Rehabilitation Act of 1973, as amended, and the regulations at 41 CFR Parts 60-1 through 60-60, 60-250, and 60-741. The affirmative action clause and regulations contained in the preceding sentence shall be incorporated by reference

# SUMMARY

This report provides a primer on the dominant caption formats and describes how Adobe Primetime enhances closed captioning workflows.

# TABLE OF CONTENTS

Introduction .....	2
Overview .....	3
Characteristics of Caption Formats .....	3
Analog Captions .....	3
Digital Captions .....	4
Online Captions .....	5
U.S. Regulations .....	7
Other Geographical Areas .....	8
Caption Formats .....	10
Teletext .....	11
EBU-STL .....	12
608 .....	12
708 .....	13
DVB .....	15
WebVTT .....	16
TTML .....	17
SMPTE-TT .....	17
EBU-TT .....	19
EBU-TT-D .....	19
SDP-US .....	20
CFF-TT .....	20
IMSC .....	21
Adobe Primetime & Closed Captions .....	22
References .....	23
About the Author .....	24

# INTRODUCTION

Over 40 years ago, the first captions appeared in analog television broadcasts, giving the hard of hearing access to the spoken words of television content [\[1\]](#). Since then, captions have become ubiquitous across digital television and are a growing presence in online video. In the modern broadcast industry, understanding and properly implementing captions is vital, both from a public welfare standpoint as well as a business and regulatory standpoint.

Although transmitting and displaying captions is a theoretically simple task, the actual technical landscape of captions is complex. Multiple languages and regulatory policies across national boundaries have led to the development of multiple caption formats. Differences that began in the time of analog broadcasts continued into the digital television era and impact the online captioning formats of today.

This report provides a primer on the dominant caption formats and describes how Adobe Primetime™ enhances closed captioning workflows.

## Scope and Terminology

In the United States, the term "closed captions" typically refers to text transcriptions that are intended for the hard of hearing [\[1\]](#). The term "subtitles" typically refers to text transcriptions that are intended for other purposes. In Europe, the term "subtitle" refers to both types of text transcriptions. The term "subtitles for the hard of hearing" is typically used to reference transcriptions that are intended for the hard of hearing. This report uses the term *captions* or *closed captions* to refer to optional captions for the hard of hearing.

A single caption format often has several names. This report uses the most common, non-ambiguous name to refer to each format.

This report focuses on captions for the hard of hearing. Other types of subtitles and accessibility features exist but will not be discussed.

This report focuses on the North American and European geographic areas. See ["Other Geographic Areas"](#).

# OVERVIEW

This section provides a high-level overview of caption formats.

## Characteristics of Caption Formats

In general, caption formats can be characterized according to certain attributes:

- **Era:** Caption formats originated during different time periods. Formats can be classified as *analog*, *digital*, or *online formats*. These categories refer respectively to formats created in the analog television era, the digital television era, and the current era of online (internet) video.
- **Use Case:** Caption formats differ by their intended use cases. *Delivery formats* are designed to be received and decoded by the end player (television, web browser, mobile device). *Authoring formats* are formats designed for export by caption authoring tools. *Interchange formats* are intended for exchange between content producers and content distributors or as a mezzanine format within the distribution system. The terms "distribution format" and "archival format" are commonly used in literature, but are avoided in this report due to their ambiguity.
- **In-band or out-of-band:** In-band caption formats are transmitted within the same stream as audio/video content. Out-of-band caption formats are transmitted in a separate stream or as a separate file (side-car).
- **Character-based or image-based:** Character-based caption formats represent captions as character codes (such as UTF-8). Image-based formats represent captions as images (typically bitmaps). Some formats can function in both character-based and image-based modes.
- **Carrier formats:** Some caption formats, especially those of the digital era, have the ability to embed data from other formats. Such a format is said to *carry* the embedded data.

## Analog Captions

This section summarizes caption formats from the analog broadcast era.

Even in today's digital and online environments, analog captions remain relevant for the following reasons:

- **Compliance:** Some nations (notably the United States) require analog captions to be embedded\* within digital transmissions.
- **Legacy:** A vast library of content exists that only contains analog captions.
- **Reach:** Analog captions reach the widest audience. By embedding analog captions within digital transmissions, both analog televisions and digital televisions can be targeted.

\* Despite their name, the dominant analog formats specify a means to translate the original analog signal into a bit stream, then specify a format for the bit stream. This allows for their carriage in digital formats.

Analog captions are primarily in one of three formats, 608, Teletext, or EBU-STL. The following table summarizes the basic characteristics of these formats:

Format	Geo	Summary
<a href="#">608</a>	US	Character-based delivery/authoring
<a href="#">Teletext</a>	Europe	Delivery/authoring Basic: Character-based Enhanced: Character-based or image-based
<a href="#">EBU-STL</a>	Europe	Character-based, authoring/interchange

## Digital Captions

This section describes caption formats used in digital television.

Digital captions are primarily in one of three formats: 708, DVB Teletext, or DVB Subtitles. The following table summarizes the basic characteristics of these formats:

Format	Geo	Summary
<a href="#">DVB Teletext</a>	Europe	Carrier for Teletext
<a href="#">DVB Subtitles</a>	Europe	Image-based or character-based delivery
<a href="#">708</a>	US	608 over 708: Carrier for 608 OR Native 708: Character-based delivery

## Online Captions

This section describes caption formats used in online video.

### Summary

Online captions are typically in one of three formats: TTML, WebVTT, or 608 over 708. The following table compares the formats:

Description	<a href="#">WebVTT</a>	<a href="#">TTML</a>	<a href="#">608 over 708</a>
<b>Actual Use Case</b> What is the format actually used for?	Delivery	Authoring/Interchange	U.S. Delivery Alternative
<b>Implementations</b>	See <a href="#">"Implementations."</a>		
<b>Industry Support</b> Which industry favors this format?	Tech	Broadcast	-
<b>Geo</b> Is the format specific to a geographic area?	Worldwide	Worldwide	U.S.
<b>Complexity</b> What is the commonly perceived complexity of the format?	Simple	Complex	-
<b>Intended Use Cases</b> What use cases were considered in the design?	Delivery	Delivery, Authoring, and Interchange	-
<b>Fragmentation</b> Are there many different variants of the format?  OR Is there a single authoritative kind of the format?	Single, agreed-upon specification	Very fragmented, many profiles	-
<b>Friendly Formats</b>	See <a href="#">"Friendly Standards."</a>		
<b>Regulations</b>	-	<a href="#">FCC Safe Harbor</a>	-

## Implementations

Historically, TTML and WebVTT have competed to become the dominant format for caption delivery. WebVTT has emerged as dominant due to its presence on all major platforms.

Official data is not available, but TTML appears to have growing support among authoring use cases and limited, but growing support among encoders as an interchange format.

The following table summarizes current delivery support:

Format	Native Players		Browsers (HTML5)			
	iOS	Android	IE	Firefox	Chrome	Safari
608 over 708	X	X				X
TTML			X			
WebVTT	X	X	X	X	X	X

*An "X" under "Native Players" indicates that the format is supported by the native media player API on the associated platform. An "X" under "Browsers" indicates that the format is supported by the HTML5 text track of the associated browser.*

## Friendly Standards

The following table summarizes relationship of online caption formats to various other formats/protocols:

Format/Protocol	Required Caption Format	Optional Caption Formats
HTTP Live Streaming (HLS)	None	608 over 708 WebVTT
DASH-264	SMPTE-TT (a profile of TTML)	608 over 708 WebVTT
HTTP Dynamic Streaming (HDS)	None	608 over 708
Interoperable Master Format (IMF)	TBD, leading candidate is IMSC (a profile of TTML)	TBD
HbbTV	EBU-TT (a profile of TTML)	



## U.S. Regulations

Recent U.S. regulations were a significant driver for online caption efforts.

In 2010, the Twenty-First Century Communications and Video Accessibility Act (CVAA) was passed in the U.S with the goal of improving the accessibility of online video. In CFR 47 § 79.4, the Federal Communications Commission (FCC) interprets the legislation into a set of specific requirements for CVAA compliance.

The FCC requirements state that online video must have captions if the content also appears on U.S. television. Since most U.S. premium online content also appears on U.S. television, the majority of U.S. premium online content must have captions in order to be in compliance. To ease the transition, the requirements have been phased in over several deadlines. The full-length live and VoD deadlines have expired and captions on this class of content are required. **The next set of deadlines extends the captioning requirements to online video clips and will take effect on January 1, 2016.**

In addition to requiring the presence of captions, the FCC requires a minimum quality level for captions. The requirements state that online captions must be provided "with at least the same quality as the television captions provided for the same programming<sup>[2]</sup>." The requirements also specifically include the ability to customize captions, but care should be taken not to emphasize this section at the expense the "same quality" requirement.

**The FCC does not require a specific format for captions, provided that they meet minimum quality levels.** During the rule review period, the FCC explicitly stated that its goal is not to mandate a single captioning format, but to let a dominant format emerge from the industry itself<sup>[3]</sup>. However, SMPTE-TT is given special treatment via the Safe Harbor Clause. Captions delivered via SMPTE-TT are automatically deemed compliant with the regulations.

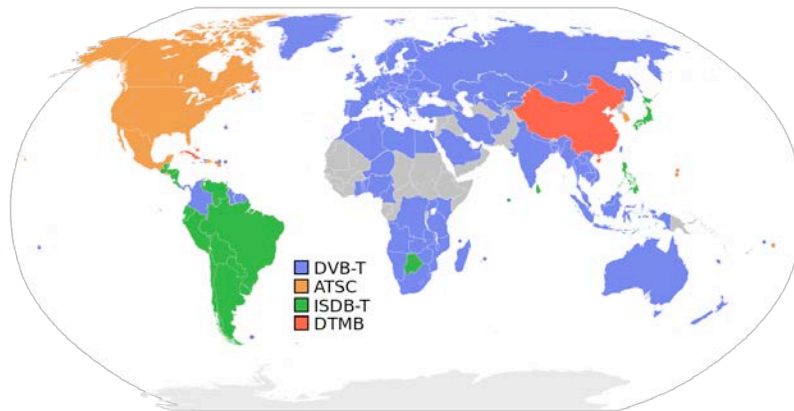
### Additional Reading:

- [Electronic Code of Federal Regulations Title 47](#) (CFR 47) describes the FCC requirements governing video accessibility. Sections 79.4 and 79.103 describe requirements that apply to online video.
- [FCC 11-138](#) is a call for comment on online caption requirements and aids an understanding of the intention of regulations.
- [FCC CVAA Homepage](#)

## Other Geographical Areas

Although not the focus of this report, this section briefly summarizes caption formats outside the United States and Europe.

Broadcast caption formats are primarily determined by the digital transmission standard that is in use. The following image from the Digital Broadcasting Experts Group summarizes the prevalence of various digital transmission standards:



Source: [http://commons.wikimedia.org/wiki/File:Digital\\_broadcast\\_standards.svg](http://commons.wikimedia.org/wiki/File:Digital_broadcast_standards.svg) <sup>[6]</sup>

**Integrated Services Digital Broadcasting (ISDB-T)** is a digital transmission standard developed and used by Japan. Character-based or image-based captions are embedded in the M2TS stream. The Association of Radio Industries and Businesses (ARIB) organization maintains and develops ISDB-T.

**ISDB-T International**, also known as *ISDB-Tb* or *SBTV-D*, is a standard derived from ISDB-T to accommodate broader use of the standard. ISDB-Tb was initially developed by Brazil, and is primarily used in South America. Literature suggests that ISDB-Tb uses the ISDB-T caption format with a modified character set, but for full confirmation a copy of the formal specification should be purchased.

South Korea uses ATSC for digital transmissions, but uses **KS C 5601** for caption delivery. KS C 5601 is a variant of CEA-708 that provides a Korean character encoding <sup>[5]</sup>.

China and several other countries use the **Digital Terrestrial Multimedia Broadcast (DTMB)** standard.

### Additional Reading

- <http://en.dtvstatus.net/>
- [http://www.arib.or.jp/english/html/overview/sb\\_ej.html](http://www.arib.or.jp/english/html/overview/sb_ej.html) (ISDB-T standards)
- <http://www.dibeg.org/techp/aribstd/harmonization.html> (ISDB-T/ISDB-Tb harmonization documents)
- <http://cpcweb.com/blog/2013/06/closed-captioning-for-south-korean-broadcast-tv/>
- <http://en.wikipedia.org/wiki/ISDB#ISDB-T>
- [http://en.wikipedia.org/wiki/ISDB-T\\_International](http://en.wikipedia.org/wiki/ISDB-T_International)
- [http://en.wikipedia.org/wiki/Digital\\_Multimedia\\_Broadcasting](http://en.wikipedia.org/wiki/Digital_Multimedia_Broadcasting)

PAL (Europe)

NTSC (US)

# CAPTION FORMATS

Teletext

608

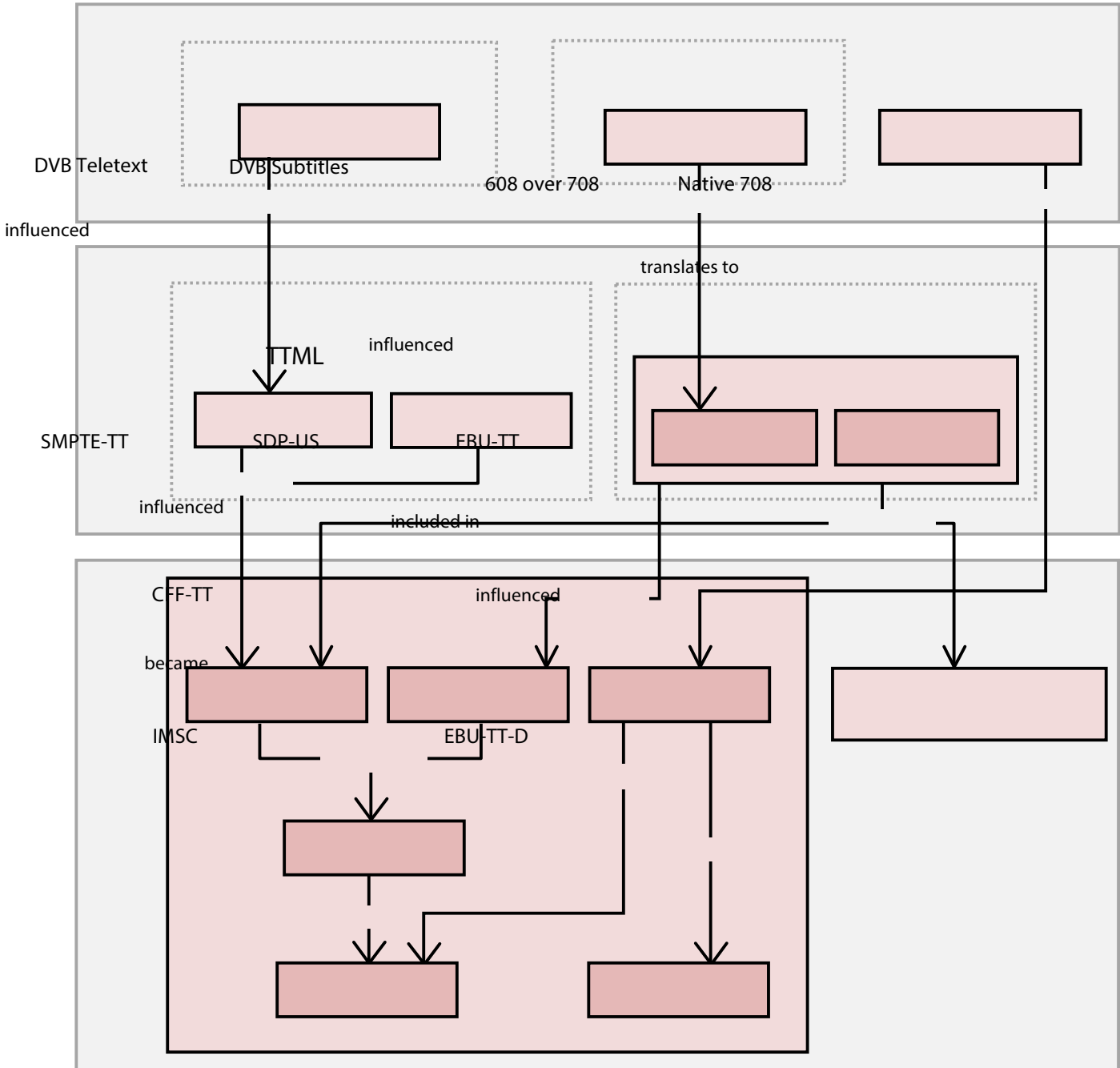
EBU-STL

carried by

This section describes each of the major caption formats. The following diagram depicts the relationships of various caption formats:

carried by

translates to



## Teletext

*Officially known as CCIR Teletext System B, ITU.R System B, CCIR 653, ITU-R BT.653, or ETSI 300 706. Also known as World System Telecom, Ceefax, or ORACLE.*

Teletext is the dominant legacy format for in-band captions in analog PAL transmissions (Europe). Teletext continues to have a presence in the digital television and online video via DVB Teletext and EBU-TT.

In analog transmissions, Teletext is carried on line 21 in the vertical blanking interval.

### Aside: Frames and the Vertical Blanking Interval

Analog television is transmitted as a series of frames, each of which has two fields. A field in an analog broadcast is analogous to a frame of modern video. A frame in an analog broadcast is analogous to a pair of frames of modern video. The signal for each field consists of series of lines. Most lines hold image data, but some are not displayed. The non-displayed lines take place during a time period known as the Vertical Blanking Interval (VBI). In analog broadcasts, the VBI lines were eventually re-purposed to transmit metadata about the broadcast, such as closed captions.

**Additional Reading:** <http://lurkertech.com/lg/fields/>

Teletext support is offered in 4 "levels", numbered 1, 1.5, 2.5, and 3.5.

Levels 1 and 1.5 are known as "Basic Teletext". Basic Teletext allows the display of captions from a fixed character set. Captions are streamed as a sequence of "pages," each of which is a set of up to 25 rows of 40 characters. Each character is represented by a code defined by the specification. Level 1.5 augments Level 1 by defining additional character codes and adding the ability to modify the current page (to allow for backward compatible transmissions). Basic Teletext allows for simple formatting, such as the customization of the text background and foreground colors.

Levels 2.5 and 3.5 are known as "Enhanced Teletext." Enhanced teletext offers significant new capabilities including: the ability to dynamically define additional characters (by specifying pixel maps over the wire), additional formatting options, extended wide-screen space, and a complex object referencing scheme.

Specifications and literature on Teletext may confuse the reader due to their mention of an obsolete use case. Teletext was originally designed to enable captioning use cases as well as a use case where interlinked informational text pages could be broadcast to a large number of televisions and terminals. The page broadcast functionality was widely used prior to the advent of the Internet, but has since become obsolete.

### Related Formats

- [DVB Teletext](#) defines a means to embed Teletext in digital transmissions.
- [EBU-STL](#) defines an archival format that may optionally use the Teletext character set.

## Additional Reading

- [ETSI 300 706](#) is the primary specification for Teletext.
- [ITU-R BT.653](#) (formerly CCIR 653) is a frequently referenced document, but is not practically useful. For historical reasons, the document provides an overview of 4 different captioning/videotex systems. The second system, System B, ultimately became dominant and is the format commonly known today as Teletext.
- The [Teletext Wikipedia page](#) is helpful for understanding the history of Teletext, but the reader should be aware that the included technical description of the format is inaccurate.

## EBU-STL

*Officially known as EBU 3264. Also known as EBU Subtitles or EBU Teletext.*

The EBU Subtitling Data Exchange Format (EBU-STL) is a legacy, out-of-band caption file format used primarily in Europe for authoring and interchange. Official usage data is unavailable, but most caption authoring tools support EBU-STL. EBU-STL files typically have the extension ".STL."

The format describes captions using a fixed character set. The format stores a sequence of Text and Timing Information (TTI) blocks. Each TTI block specifies a row of characters, an associated time code, and other metadata. The character codes used for the captions may optionally match the Teletext codes.

### Related Formats

- [Teletext](#) characters may be embedded within an EBU-STL file.
- [EBU-TT](#) is the online successor to EBU-STL.

## Additional Reading

- [EBU 3264](#) is the specification for EBU-STL.

## 608

*Officially known as CEA-608 or EIA-608 . Also known as 608/708 (ambiguous) or Line 21 Captions.*

608 is the primary format for in-band captions in analog NTSC transmissions (US). 608 continues to have use in both digital and online video through 608 over 708.

In the United States, the FCC mandates that analog captions be in the 608 format. In analog transmissions, 608 is carried in the [vertical blanking interval](#).

The majority of premium content produced for the United States today still contains 608 captions embedded in the [608 over 708](#) digital format. Despite the analog to digital transition, new digital content in the U.S. often contains 608 captions to accommodate (1) legacy television sets that utilize digital to analog converters and (2) existing production workflows. Legacy content also contain 608 captions.

608 is a streaming, character-based format that allows for the transmission of up to 4 simultaneous channels of data (CC1, CC2, CC3, CC4). Data is transmitted in 2 bytes chunks, each of which represent either 2 characters or a two-byte command (PAC). PACs control various display characteristics such as display position, formatting, and animation style. Captions may be animated in a number of ways: Roll-up, Pop-on, and Paint-on.

### Related Formats

- 608 can be carried within [708](#).
- 608 can be translated to or tunneled through [SMPTE-TT](#), [CFF-TT](#), and [IMSC](#).
- 608 can be translated to [WebVTT](#).
- 608 heavily influenced the design of most [TTML](#) profiles.

### Additional Reading

- [47 CFR § 79.101](#) (formerly [47 CFR § 15.119](#)) of the FCC guidelines serve as the primary specification of the 608 format.
- [CEA-608](#) (formerly EIA-608) provides enhancements and additional requirements for 608 implementations.

## 708

*Officially known as CEA-708. Also known as 608/708 (ambiguous).*

708 is the primary format for in-band captions in digital ATSC transmissions. ATSC is the current standard for digital terrestrial transmission in the United States. The [608 over 708](#) portion of the format is one of the online caption delivery formats in use today.

708 is a streaming format that can carry two types of data simultaneously: "[608 over 708](#)" and "[native 708](#)". The format achieves this through a layered design where the lowest layer of the format, the DTVCC Transport Layer, multiplexes 608 and native 708 data.

In ATSC and online video, 708 captions are carried within the MPEG-2 TS stream as an H.264 SEI NALU.

### 608 over 708

*No official name. Also known as legacy 608 or compatibility bytes.*

A 708 stream may carry 608 data via a practice commonly known as 608 over 708. CEA-708 requires that 708 decoders must also be able to decode embedded 608 captions. Digital to analog TV converters insert the embedded 608 into the vertical blanking interval of the analog signal.

For regulatory conformance, 608 over 708 data is always present in US digital transmissions <sup>[4]</sup>.

The iOS, Android, and Safari video players are able to decode the 608 portion of 708.

## Native 708

*No official name. Also known as DTVCC captions or true 708.*

A 708 stream may also carry caption data in native 708 format. Although the native 708 format has similarities with the 608 format, the two formats differ enough that they should be considered distinct.

The native 708 feature set is a superset of the 608 feature set. Notable features introduced by native 708 include a windowing system, transparency, variable font size, and user customizable attributes such as font size.

For regulatory conformance, native 708 is always present in US digital transmissions<sup>[4]</sup>. Since 608 over 708 is also required, captions are typically authored in 608 and converted to 708 by encoders or authoring tools. Such captions are known as upconverted, translated, or transcoded captions<sup>[4]</sup>.

No online video player officially supports native 708.

## Related Formats

- [608](#) can be carried within 708.
- 708 can be translated to or tunneled through [SMPTE-TT](#), [CFE-TT](#), and [IMSC](#).
- 708 can be translated to [WebVTT](#).
- 708 heavily influenced the design of most [TTML](#) profiles.

## Additional Reading

- [CEA-708](#) is the primary specification for 708. Section 4.3 describes 608 over 708 and the DTVCC transport layer. Sections 5-8 describe Native 708 captions (referred to as the DTVCC Interpretation, Coding, Service, and Packet Layers).
- [SCTE 128](#) Section 8.2 specifies how 708 data should be embedded in an ATSC MPEG-2 TS stream. [A/72 Part 2](#) Section 6.2 and [A/53 Part 3](#) describe the same process in a less concise manner.
- [47 CFR § 79.102](#) describe FCC rules regarding digital television captions. These rules define additional technical requirements for the implementation of CEA-708. Notably, the rules sometimes specifying new functionality such as additional columns for 608 over 708 captions when displayed on 16:9 screens.
- "[Implementing Closed Captioning for DTV](#)" describes implications of digital television FCC regulations.
- [Apple HTTP Live Streaming Documentation](#) describes 608 over 708 carriage in HLS.
- [Android MediaFormat Documentation](#) mentions 608 over 708 as a caption format.



## DVB

DVB transmissions may embed captions in two possible formats:

- [DVB Teletext](#)
- [DVB Subtitles](#)

Official data is scarce, but most sources indicate that DVB Teletext is more prevalent in Europe than DVB Subtitles.

### DVB Teletext

*Officially known as ETSI 301 775 or ETSI 300 472 . Also known as DVB-TXT or DVB-VBI.*

DVB Teletext embeds Teletext in a DVB MPEG-2 TS stream as a PES packet of type 6.

#### Related Formats

- [Teletext](#)

#### Additional Reading

- [ETSI 301 775](#) (formerly [ETSI 300 472](#)) specifies DVB Teletext, as well as a method for embedding other VBI data in a DVB MPEG-2 TS Stream.
- [ETSI 300 468](#) defines a means for signalling the presence of DVB Teletext on an DVB MPEG-2 TS stream.

### DVB Subtitles

*Officially known as ETSI*

DVB Subtitles allows either image-based or character-based captions to be embedded within a DVB MPEG-2 TS stream. Image-based captions appear to dominate use of the format. DVB Subtitles are intended for use by non-European languages.

DVB subtitles are transmitted as PES packets of type 6. In the format, a PES packet may contain multiple segments. An object data segment carries the data for an object, which is either image or character data. Composition segments position objects on the screen and are displayed at their presentation times. Image data is expressed as run length encoded, indexed pixel values. The character data is expressed as a series of byte codes, but notably the meaning of the codes is left undefined. The specification suggests that "a local agreement between broadcasters and equipment manufacturers may be an appropriate way to ensure reliable operation of character coded subtitles."

#### Additional Reading

- [ETSI 300 743](#) specifies DVB Subtitles.
- [ETSI 300 468](#) defines a means for signalling the presence of DVB Teletext on an DVB MPEG-2 TS stream.

## WebVTT

*Formerly known as WebSRT.*

The Web Video Text Tracks Format (WebVTT) is a format for delivery of internet captions. A WebVTT file contains the following:

- **Cues:** A WebVTT file consists primarily of a sequence of cues. A cue is one or more lines of text with an associated time interval and optional metadata. Styles may be applied to the cue text using a simple markup syntax. Portions of the cue text may also be designated to appear at particular times in order to achieve paint-on style animations.
- **Regions:** A cue may also be associated with a region, which specifies the bounding box of the text when rendered on screen.
- **Header:** A WebVTT file begins with a header that includes document-scoped metadata.

WebVTT integration with HTML5 is well defined. The specification defines the VTT Cue and VTT Region APIs to access WebVTT-specific information from the TextTrackCue API. The specification also defines a method for CSS to further customize the appearance of cues.

WebVTT is a W3C Draft Community Group Report and not presently on the W3C Standards Track.

WebVTT was originally developed by Google.

### See also

- [Online Captions](#)
- [Design Analysis: TTML vs. WebVTT](#)

### Related Formats

- [608](#) can be converted to WebVTT.
- [708](#) can be converted to WebVTT.
- HLS specifies semantics for associating WebVTT with a presentation.

### Additional Reading

- [Official Specification](#)
- "[Conversion of 608/708 Captions to WebVTT](#)" defines a method for generating WebVTT from 708.

# TTML

Formerly known as *Distribution Format eXchange Profile (DXFP)*.

The Timed Text Markup Language (TTML) is an XML-based format for delivery, interchange, and authoring of internet captions. TTML exposes multiple types of elements and attributes:

- **Content elements** are HTML-like elements (such as <div>, <p>, <span>) that contain the caption text.
- **Timing attributes** specify the time interval during which content should be visible. Timing attributes may also be applied to layout and animation elements.
- **Style elements** specify the appearance of content via a simple XML-based styling system.
- **Layout elements** specify the layout properties (such as the bounding boxes) of content.
- **Animation elements** can be used to alter the style of text at particular times.
- **Metadata elements** specify additional metadata about the presentation.

TTML allows for the definition of *profiles*, which can be loosely thought of as variants of TTML. The base TTML specification defines a broad set of features and the xml semantics for how a TTML document will express those features. Future specifications are expected to define *profiles*, each of which is a set of features, extensions (new features), and requirements to ensure interoperability.

The base TTML specification defines a few profiles, the most significant of which is the *DXFP Full Profile*. The full profile includes all features of TTML.

TTML is a W3C Recommendation.

## See also

- [Online Captions](#)
- [Design Analysis: TTML vs. WebVTT](#)

## Related Formats

- [SMPTE-TT](#), [EBU-TT](#), [EBU-TT-D](#), [IMSC](#), [SDP-US](#), and [CFF-TT](#) are profiles of TTML.

## Additional Reading

- [Official Specification](#)

## SMPTE-TT

The SMPTE Timed Text Format (SMPTE-TT) is a profile of TTML developed primarily by the broadcast industry. SMPTE-TT is notable for its mention in the FCC Safe Harbor clause and its consideration of legacy formats.

SMPTE-TT requires all features of TTML.

SMPTE-TT defines extensions to aid the transition from legacy analog and digital formats. The ability to translate 608, 708, DVB Subtitles, and Teletext captions to SMPTE-TT documents are explicitly mentioned as explicit design goals. EBU-STL is not explicitly mentioned. SMPTE-TT introduces multiple features to aid translation:

- **Tunneling:** The original analog/digital captions may be embedded as binary blobs. Blobs may be associated with the document as a whole or with a particular time interval.
- **Images:** Div tags may have background images in order to aid translation from image-based formats (DVB Subtitles).
- **Additional Metadata:** Information about how a document was translated may be embedded in the document header. A document may specify an origin format (as a URI) and the translation's fidelity level ("Preserve" or "Enhance"). Together this information allows for a SMPTE-TT document to state that it was translated according to a specific specification.

The SMPTE-TT specification belongs to a family of documents collectively known as SMPTE 2052. The 2052 family also includes recommended practices (RP's) on how to translate specific legacy formats to SMPTE-TT. Although these documents are not official standards, they define extensions, specify the byte format for tunneled data, and more precisely define Preserve and Enhanced translation modes for a particular origin format. Therefore, the RP's should be considered a part of "implementing SMPTE-TT." Currently, RP's have been published that describe how to translate 608 to SMPTE-TT and 708 to SMPTE-TT.

#### **Additional Reading**

- [SMPTE 2052-1](#) defines SMPTE-TT.
- [RP 2052-10](#) describes how to convert 608 to SMPTE-TT.
- [RP 2052-11](#) describes how to convert 708 to SMPTE-TT.

#### **Related Formats**

- [TTML](#)
- SMPTE-TT is the official caption format of DASH-264.
- SMPTE-TT can be generated from [608](#) and [708](#).
- [IMSC](#) and [CFF-TT](#) incorporate SMPTE-TT extensions.

## EBU-TT

*Also known as EBU Tech 3350.*

EBU Timed Text (EBU-TT) is a profile of TTML developed as a successor to EBU-STL. The intended use case for EBU-TT is as an authoring and interchange format.

EBU-TT differs from the DXFP Full profile in the following ways:

- **Tunneling:** A binary blob may be embedded within the document. Unlike SMPTE-TT tunneling, an EBU-TT binary blob may only be associated with the document as a whole.
- **Well-defined mapping from EBU-STL:** EBU 3360 specifies how EBU-STL documents should be translated to EBU-TT documents.
- **Reduced Required Features:** Animation elements are not required as are various style/timing/layout options.
- **Additional Metadata:** Program, episode, copyright and other information may be embedded in the document header.

### Additional Reading

- [EBU Tech 3350](#) is the official specification of EBU-TT.
- The [EBU-TT Homepage](#) describes the EBU-TT family of standards.
- [EBU Tech 3360](#) specifies how EBU-STL documents should be translated to EBU-TT documents.

### Related Formats

- [TTML](#)
- EBU-TT is the successor to [EBU-STL](#)
- [EBU-TT-D](#)

## EBU-TT-D

*Also known as EBU Tech 3380.*

EBU-TT-D is a related, but distinct profile from EBU-TT. Unlike EBU-TT, EBU-TT-D's intended use case is as a delivery format. EBU-TT-D shares many characteristics with EBU-TT, but lacks tunneling and metadata features.

### Additional Reading

- [EBU Tech 3380](#) is the official specification of EBU-TT-D.
- The [EBU-TT Homepage](#) describes the EBU-TT family of standards.
- [EBU Tech 3381](#) specifies how to embed EBU-TT-D within ISOBMFF, with DASH as the intended use case.

### Related Formats

- [TTML](#)
- [EBU-TT](#)
- EBU-TT-D is the official caption format of HbbTV.

## SDP-US

*Also known as SDP.*

The Simple Online Delivery Profile (SDP-US) is a profile of TTML that reduces the required feature set to those deemed necessary for CVAA compliance. As such, SDP-US focuses on 608/708 and notably provides analysis of how TTML features map to CVAA compliance. The resulting feature set is a moderately-sized subset of DXFP Full Profile. SDP-US is not known to have significant adoption. SDP-US was originally proposed by Microsoft, but was published by the W3C as a Working Group Note.

### Additional Reading

- [Official Specification](#)

### Related Formats

- [TTML](#)
- [IMSC](#)

## CFF-TT

Common File Format Timed Text (CFF-TT) is a pair of TTML profiles by UltraViolet. CFF-TT differs from the DXFP Full Profile in the following ways:

- **Separate Image and Text Profiles:** CFF-TT introduces two profiles that represent either character-based captions or image-based captions. Both may not be used simulatenously.
- **SMPTE-TT extensions:** CFF-TT incorporates SMPTE-TT extensions.
- **Performance constraints:** CFF-TT explicitly introduces limits on the complexity of documents. This implicitly introduces minimum (and maximum) performance expectations for decoders.
- **Reduced required feature set**

### Additional Reading

- [Official Specification](#)

### Related Formats

- [TTML](#)
- [IMSC](#)

## IMSC

Text and Image Profiles for Internet Media Subtitles and Captions (IMSC) is a pair of TTML profiles developed primarily as a candidate format for the Interoperable Media Format effort.

### Aside: Interoperable Media Format (IMF)

Interoperable Media Format (IMF) is a SMPTE effort to develop a single mezzanine media file format that can be used as a "grand master" interchange format. IMF was originally intended for the transfer of media from studios to distributors, but its scope has expanded to archival use also. Netflix is a notable supporter of IMF. IMF is not currently intended as the format for internet delivery to end-clients.

**Additional Reading:** [Video Presentation by Editor](#)

IMSC is heavily derived from CFF-TT, but incorporates aspects of EBU-TT. IMSC-TT is a W3 Candidate Recommendation.

### Additional Reading

- [Official Specification](#)

### Related Formats

- [TTML](#)
- [CFF-TT](#)
- [EBU-TT](#)

## ADOBE PRIMETIME & CLOSED CAPTIONS

Adobe is widely recognized by the industry as a leader in video accessibility. In 2014, the FCC awarded Adobe Primetime the prestigious FCC Award for Advancement in Accessibility (Chairman's AAA).

Adobe understands that closed captions are a complex, yet vital part of the premium video ecosystem. To address your closed captioning needs, Adobe built robust closed captioning support as a key component of the Primetime TVSDK. The TVSDK provides a best-in-class closed captioning implementation that enables 608 over 708 and WebVTT captions across your target platforms.

The TVSDK extends the reach of premium content by bringing industry standard format support to the following platforms:

Format	Primetime TVSDK					
	iOS	Android	Xbox One	Roku	Browser (Flash)	Browser (HTML5 MSE)
608 over 708	X	X	X	X	X	
WebVTT	X	X	X	X	X	X

The TVSDK makes use of industry standard practices when packaging captions with content. 608 over 708 captions are embedded in MPEG-2 TS streams according to SCTE 128, and WebVTT captions can be associated with an HLS playlist according to the standard practices defined by the HTTP Live Streaming specification.



## REFERENCES

1. [http://en.wikipedia.org/wiki/Closed\\_captioning](http://en.wikipedia.org/wiki/Closed_captioning)
2. <http://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=8b6798f33bd3185de2583e513a9b6ba4&r=PART&n=pt47.4.79>
3. [https://apps.fcc.gov/edocs\\_public/attachmatch/FCC-11-138A1.pdf](https://apps.fcc.gov/edocs_public/attachmatch/FCC-11-138A1.pdf)
4. <http://www.dcmp.org/caai/nadh219.pdf>
5. <http://cpcweb.com/blog/2013/06/closed-captioning-for-south-korean-broadcast-tv/>
6. [http://commons.wikimedia.org/wiki/File:Digital\\_broadcast\\_standards.svg](http://commons.wikimedia.org/wiki/File:Digital_broadcast_standards.svg)

## **ABOUT THE AUTHOR**

Glenn Eguchi is a Senior Computer Scientist in the Adobe Primetime Video Solutions Architecture Team. When he is not thinking about closed captions, Glenn likes to spend his time making games, playing guitar, and staying active.